

Amendments to the Specification:

Please amend paragraph [0079] as follows:

A percent identity for any subject nucleic acid or amino acid sequence (e.g., any of the hydroxylase polypeptides described herein) relative to another “target” nucleic acid or amino acid sequence can be determined as follows. First, a target nucleic acid or amino acid sequence can be compared and aligned to a subject nucleic acid or amino acid sequence, using the BLAST 2 Sequences (Bl2seq) program from the stand-alone version of BLASTZ containing BLASTN and BLASTP (e.g., version 2.0.14). The stand-alone version of BLASTZ can be obtained at fr.com/blast or ncbi.nlm.nih.gov on the World Wide Web <www.fr.com/blast> or <www.ncbi.nlm.nih.gov>. Instructions explaining how to use BLASTZ, and specifically the Bl2seq program, can be found in the ‘readme’ file accompanying BLASTZ. The programs also are described in detail by Karlin et al. (1990) *Proc. Natl. Acad. Sci.* 87:2264; Karlin et al. (1990) *Proc. Natl. Acad. Sci.* 90:5873; and Altschul et al. (1997) *Nucl. Acids Res.* 25:3389.

Please amend paragraph [0083] as follows:

The identification of conserved regions in a template, or subject, polypeptide can facilitate homologous polypeptide sequence analysis. Conserved regions can be identified by locating a region within the primary amino acid sequence of a template polypeptide that is a repeated sequence, forms some secondary structure (e.g., helices and beta sheets), establishes positively or negatively charged domains, or represents a protein motif or domain. See, e.g., the Pfam web site describing consensus sequences for a variety of protein motifs and domains at sanger.ac.uk/Pfam and genome.wustl.edu/pfam/ on the World Wide Web <http://www.sanger.ac.uk/Pfam/> and <http://genome.wustl.edu/Pfam/>. A description of the information included at the Pfam database is described in Sonnhammer et al. (1998) *Nucl. Acids Res.* 26: 320-322; Sonnhammer et al. (1997) *Proteins* 28:405-420; and Bateman et al. (1999) *Nucl. Acids Res.* 27:260-262. From the Pfam database, consensus sequences of protein motifs and domains can be aligned with the template polypeptide sequence to determine conserved region(s).

Please amend paragraph [0103] as follows:

Typically, a nucleic acid of the invention has 70% or greater sequence identity, e.g., 80%, 85%, 90%, 95%, 96%, 97%, 98%, or 99% or greater sequence identity to a target nucleic acid. Sequence identity is determined as described herein. In some embodiments, nucleic acids are from 20 to 30 nucleotides, or 20 to 50 nucleotides, or 25 to 100 nucleotides, or 500 to 1500 nucleotides, or 900 to 2,000 nucleotides in length. Specific embodiments of nucleic acids include nucleotide sequences set forth in the sequence listings. It is noted that the degeneracy of the genetic code permits codon modification without a corresponding modification of the amino acid sequence. Thus, codons in a nucleic acid can be modified if desired, which may optimize expression of a polypeptide. For example, codons with 8% or lower percentage occurrence in a selected plant species genome can be replaced with a more frequently occurring codon, e.g., the most frequent or second most frequent codon for that particular amino acid. As another alternative, one member of a contiguous pair of codons can be modified if both codons have an occurrence of 12% or lower in known sequences of the genome of a selected plant species. Data relating to codon usage database can be found, for example, at kazusa.or.jp/codon/ on the World Wide Web <<http://www.kazusa.or.jp/codon/>>. Codons can also be changed to remove ATTAA (i.e., AUUUA) elements which may contribute to mRNA instability, and codons may be changed to ablate potential polyadenylation sites. Codons can also be modified to break up runs of five or greater contiguous nucleotides of A, G, C or T (e.g., TTTTTT). Codons can also be modified to reduce the likelihood of aberrant splicing. Splicing potential can be assessed and donor (GT) or acceptor (AG) splice sites ablated in order to diminish splicing potential, using predictive algorithms such as algorithms at cbs.dtu.dk/services/NetPGene on the World Wide Web [[<<http://www.cbs.dtu.dk/services/NetPGene>>]]. In addition, codons near the N-terminus of the polypeptide can be changed to codons preferred by a selected plant species, e.g., soybean (*Glycine max*). It will be appreciated that one or more codon modifications, including but not limited to the modifications discussed above can be made to a nucleic acid coding sequence. Examples of sequences that have one or more codon modification(s) to improve plant expression and have slight changes to the amino acid sequences relative to the wild-type sequence include SEQ ID NOS: 28 through 33.

Please amend paragraph [0182] as follows:

Four codon optimized hydroxylase (*Ricinus communis* and *Lesquerella gracilis*) and epoxxygenase (*Stokesia laevis* and *Crepis bennies*) sequences were constructed as follows. First the 2nd, 3rd, and 4th codons downstream of the initiation methionine codon were changed to GCT, TCC, and TCC (encoding alanine, serine and serine). Secondly, codons with 8% or lower percentage occurrence in either the *Arabidopsis thaliana*, *Glycine max*, *Lycopersicon esculentum* or *Nicotiana tabacum* genomes (e.g., CGG for arginine) were replaced with the most frequent or second most frequent codon for that particular amino acid (e.g., AGA or AGG for arginine). Finally, one member of a contiguous pair of codons was optimized if both codons had an occurrence of 12% or lower in either the *Arabidopsis thaliana*, *Glycine max*, *Lycopersicon esculentum* or *Nicotiana tabacum* genomes. Data for the codon optimization process were taken from the codon usage database ([kazusa.or.jp/codon/](http://www.kazusa.or.jp/codon/) on the World Wide Web [[<http://www.kazusa.or.jp/codon/>]]).

Please amend paragraph [0183] as follows:

Codons were also changed to remove ATTTA (i.e., AUUUA) elements which may destabilize mRNAs, to ablate potential polyadenylation sites, and to break up runs of A, G, C or T of five or greater nucleotides (e.g., TTTTT). Codons were also modified to reduce the likelihood of aberrant splicing. Splicing potential was assessed with the NetPlantGene prediction server (cbs.dtu.dk/services/NetPGene/ on the World Wide Web <http://www.cbs.dtu.dk/services/NetPGene/>). Whenever a donor and acceptor existed where both were predicted with greater than 0.9 confidence a codon was mutated to ablate either the donor (GT) or acceptor (AG) sites and thus diminish splicing potential. SEQ ID NOS: 30, 31, 32 and 33 are examples of these optimized sequences.

Please amend paragraph [0184] as follows:

Additional codon optimized variants of the *Ricinus communis* hydroxylase and *Stokesia laevis* epoxgenase genes were made. These additional sequences contained modifications to more closely mimic the most common soybean (*Glycine max*) codons. The 2nd, 3rd, and 4th codons downstream of the initiation methionine codon were changed to GCT, TCC, and TCC

(encoding alanine, serine and serine). Codons were also changed to remove ATTTA (i.e., AUUUA) elements which may destabilize mRNAs, to ablate potential polyadenylation sites, and to break up runs of A, G, C or T of five or greater nucleotides (e.g., TTTTT). Codons were also modified to reduce the likelihood of aberrant splicing. Splicing potential was assessed with the NetPlantGene prediction server (cbs.dtu.dk/services/NetPGene/ on the World Wide Web <http://www.cbs.dtu.dk/services/NetPGene/>). Whenever a donor and acceptor existed where both were predicted with greater than 0.9 confidence a codon was mutated to ablate either the donor (GT) or acceptor (AG) sites and thus diminish splicing potential. Data for codon optimization procedures were taken from the codon usage database (<http://www.kazusa.or.jp/codon/>). SEQ ID NOS: 28 and 29 are examples of such optimized *R. communis* and *S. laevis* hydroxylase and epoxygenase genes, respectively.